Department of Electrical and Electronic Engineering

Imperial College London

Control of
Gene Expression

2nd Year Group Project

PROJECT SUPERVISOR

DR. IMAD JAIMOUKHA

GROUP MEMBERS

AZUBUIKE NDUKWE (ACN08)

CHARITH AMARASINGHE (CA508)

GONZALO DE GISBERT LOPEZ (GED08)

MENG SIEW (MKS08)

SAM HARRASI (SH1708)

ZHENG ZHOU (ZZ1008)

ZHIHANG LIANG (ZL408)

ABSTRACT

As a relatively young field of science, Genetic Engineering is swiftly revolutionizing medical treatments. At the heart of genetic engineering lies the concept of gene expression, found in every cell and affecting every bodily function of organisms. Synthetic biology, as the main branch of genetic engineering, is bringing new hopes for the treatment of diseases previously thought to be incurable. However, as the complexity of the genetic networks increases, the need for a powerful way of conceptualizing these networks becomes more essential. In this Literature review, a novel approach that employs control engineering to analyse, understand ,and design genetic networks is explored. A brief introduction to the process of gene expression followed by certain factors affecting it, will serve as background knowledge. The investigation will then consider initial methods employed by scientists to control the expression of genes, in particular, the production of Insulin. Recently developed methods which aim to modularize genetic networks as digital gates will then be thoroughly analyzed. This level of abstraction will help introduce the concept of feedback in genetic networks. Finally, potential future techniques for controlling gene expression will be evaluated. These methods will focus on using already existing bodily functions with the hope of creating independent gene expression control mechanisms. Though no explicit control theory formula is presented, every effort has been made to draw an analogy between genetic networks and control systems. The ultimate aim of this report is to motivate the concept, that if these technologies are successfully developed, many diseases could one day be cured.

Literature review

Key words: Synthetic biology, Recombinant DNA, Genetic circuit building blocks

TABLE OF CONTENTS

INTRODUCTION

As technological innovation continues at an unpredictable rate, new areas of research and development are emerging. Current systems continue developing, while research is progressively blurring the distinction between different areas of science and engineering. As the spread of new diseases influences the way we carry out our lifestyles, research in the fields of Biology, Engineering and Medicine is being stimulated. One topic actively being affected by this is the study of gene expression and how it can be controlled.

Gene expression describes a process through which proteins are produced for an organism to utilize. The human genome has thousands of different genes, each of which can be expressed at any particular time. Gene expression involves the transformation of an individual, or set of genes into their gene products: proteins. Each gene can be thought of as a library book, containing information on how to synthesize a protein molecule or in some instances, a non-coding RNA (Twyman 1). If it were possible to externally control the expression of genes, man-made biological systems would become a reality.

For centuries, ecosystems have inspired the field of engineering to model and simulate natural systems and processes. As a consequence, control theory has emerged attempting to correlate the inputs and outputs of systems through mathematical relationships. These relationships have enabled input changes in a system to yield predictable outputs (Doyle, Bruce, and Tannenbaum 4). There is growing interest in applying this knowledge back to its origins, in nature.

The last few decades have seen engineers and scientists investigating the possible fusion of engineering concepts and gene expression. If such an applied practice were possible, its biggest benefactor would be the medical world, potentially being able to cure any disease (UT Southwestern Medical Center 1). Many other fields would also benefit from this, due to the inherent advantages of biological systems, principally the ability to evolve (Yokobayashi, Weiss, and Arnold 16587-91).

During this investigation, the processes and mechanisms of gene expression will be explored. While there are numerous factors and variables involved regarding the operation of biological systems, a macro level approach will be taken when investigating the control of these systems. The mechanisms required for potential regulation of gene expression, particularly in single cell organisms i.e. prokaryotes, will be analysed through a study of engineered bacteria used to produce Insulin. Finally, the future of these techniques, with emphasis on the creation of genetic decision-making networks will be discussed. The social and ethical issues concerning the alteration of biological systems will also be addressed.

GENE EXPRESSION

Although heredity has always been present in living organisms, the reasons for it were not known for many years. Advancements in our understanding of biology have shown the existence of a genome, which can be thought of as an organism's blueprint, containing all the information describing it.

Information within the genome is divided into units called genes. The information within these genes is coded on strands of Deoxyribonucleic Acid (DNA).

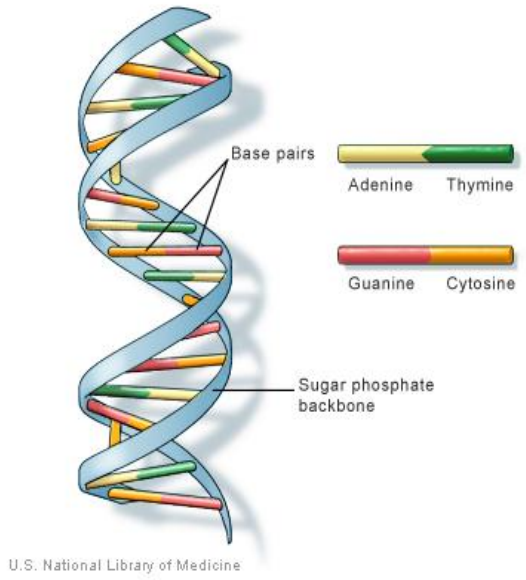


Figure 1-1 DNA Structure

DNA is composed of 4 chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Each base is connected to a sugar molecule, a phosphate molecule, and another base. This arrangement of two molecules and two bases is called a nucleotide (National Library of Medicine 1). The DNA structure is composed of recurring nucleotides, resulting in a long chain of bases.

The genetic code within DNA, is used by living organisms to synthesize types of proteins and RNA that are essential towards cellular functions and the sustainment of an organism as a whole. From the colour of an organism's skin, to the regulation of a human's heart rhythm, virtually every bodily function is controlled by gene expression (Goodchild 1).

It is widely accepted that in multi-cellular organisms, different cell types mainly have the same genetic code (Latchman 3). These cell types will only express a certain section of their genes during their lifetime. For example, skins cells will express a certain segment of genetic code, while liver cells will express another segment. The genes expressed in a cell are regulated by a number of mechanisms during different stages of the gene expression pipeline. However, in order to examine these mechanisms, the stages of gene expression must first be understood.

The process of gene expression involves two key stages: transcription and translation. In transcription, a Ribonucleic Acid (RNA) transcript is created from a DNA segment. The RNA acts as a chemical messenger which then controls the production of a protein during the translation

stage. The figure below illustrates the phases of Gene Expression. The sections below will explore each stage in further detail.

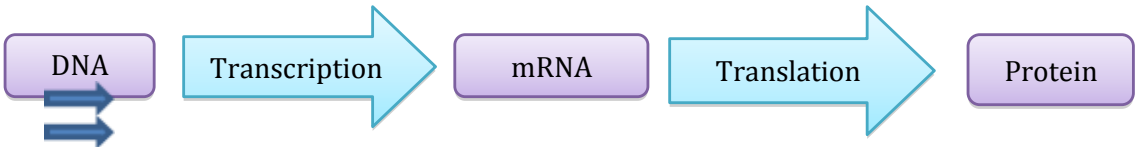


Figure 1-2 Overview of gene expression

TRANSCRIPTION

Transcription describes the process by which a segment of DNA corresponding to a particular gene is copied to create a strand of ribonucleic acid (RNA).

The majority of this task is performed inside a cell by a chemical 'machine' called RNA polymerase (RNAP). While DNA is composed of paired bases, only one set of bases is copied to create the RNA. RNAP is an enzyme that can determine where to start and stop copying DNA segments by recognising certain sequences of bases in the strand. The RNAP product is termed Messenger RNA (mRNA) since it often acts as a chemical signal to another part of the cell (Pierce 348).

TRANSLATION

The process by which the genetic code within mRNA directs the synthesis of a protein is called Translation.

Proteins are commonly composed of up to 20 different types of amino acids (Lewin 71). Variations in the combination of amino acids results in different proteins. Using the set of bases found in the mRNA, a ribosome associates a consecutive sequence of 3 bases to one type of amino acid. This combination of 3 bases per amino acid is called a Codon. The ribosome assembles a protein by reading successive units of information from the mRNA until it comes to a codon that tells it to stop (Pierce 409).

The resultant proteins play critical roles in the structure, function and maintenance of cells and whole organisms. Almost every single chemical process in an organism's body requires proteins, mainly as enzymes. For example, insulin is a protein produced by the cells in the pancreas. It functions to reduce any excessive concentrations of blood sugar.

CONTROL OF GENE EXPRESSION

As previously seen, different cells, depending on the nature of their function, only express certain genes during their lifetime. This relies on the fact that specialized cells show a large distinction in their protein and RNA content. Analyzing the gene expression process, conclusions to these variations may be formed regarding possible changes in the cell's DNA or due to some other mechanisms which result in different gene expression patterns.

Research has shown that in almost all circumstances, DNA across different cell types is practically the same (Latchman 2). This substantiates the idea that other mechanisms in a cell affect gene expression. Furthermore, assumptions can be made that these mechanisms occurs before, during, or after the processes of Transcription and Translation. However, it has been found that in Prokaryotes, this regulation occurs mostly during the Transcription phase and rarely in Translation (Biotol, Currell, and Dam-Mieras 170).

It is likely that such a mechanism acts prior to mRNA synthesis, due to energy preservation. Years of evolution have caused cells to progressively become more efficient in their processes, and this may be one of various reasons as to why regulation occurs during transcription (Biotol, Currell, and Dam-Mieras 170).

PROMOTERS

In order to transcribe a gene from DNA, RNA polymerase (RNAP) must first attach to the promoter site of that gene. A promoter site is a sequence of bases in DNA which can easily bind with a segment of the RNAP molecule. Once RNAP has bound to the gene, it will move along the DNA strand and create the RNA transcript. Some promoter sites are classified as weak or strong depending on how readily they attract RNAP, and as such, strong promoters lead to a higher rate of transcription.

REGULATION OF GENE EXPRESSION

In prokaryotes, single cell organisms such as bacteria, the presence of certain proteins can be found to inhibit or encourage the transcription of certain genes. Since the promoter site is where RNAP attaches, controlling gene expression can be achieved by regulating the ability of RNAP to bind to this promoter site. There are a number of models which can explain this behaviour (Latchman 203).

In the first case, the genes have a strong promoter site, hence are active by default. A certain type of protein, called a Repressor Protein, is introduced into the system, and binds to the promoter site to prevent RNAP from attaching to the gene and transcribing it *(See Figure 2-1a).*

Another instance involves the constant presence of the repressor in the cell; hence the gene is inactive by default. If the gene needs to be activated, the repressor can be isolated by the introduction of another protein called an Inducer Protein. The repressor will prefer to bind with the inducer rather than the promoter site since the resultant combination is more energetically favourable. This unblocks the promoter site for RNAP to begin transcription.

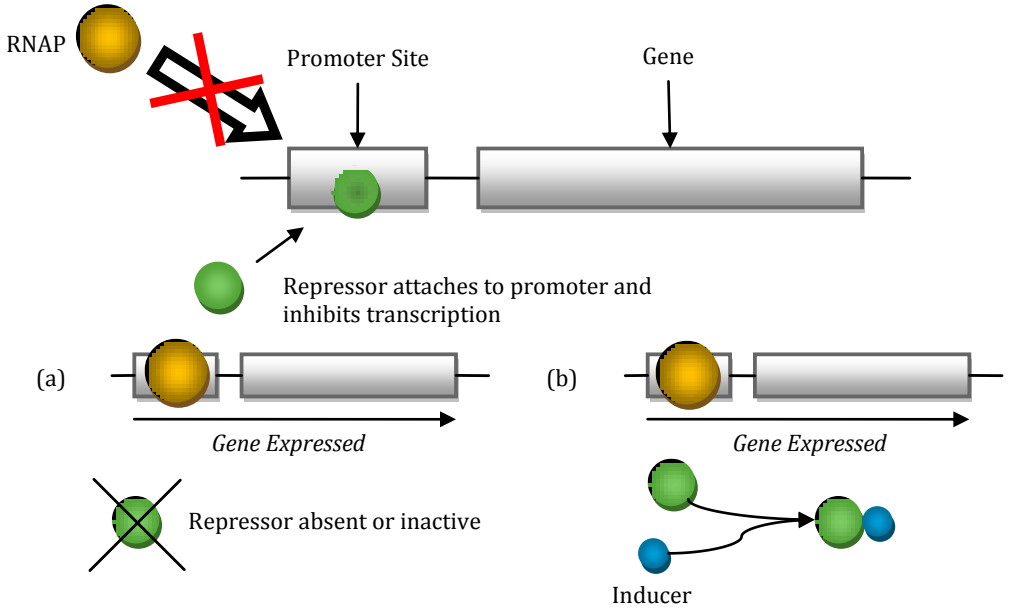


Figure 2-1 Action of a repressor

The model discussed above assumes a strong promoter site which is able to attract RNAP and transcribe to the mRNA. The contrary to this model, whereby the promoter site is weak and naturally cannot easily attract RNAP also exists. In this case, there is not a sufficient rate of transcription.

In order to help ensure the gene would be expressed, the introduction of a protein to the system, termed an Activator, is necessary. The activator would attach to the promoter site and improve its ability to attract RNAP leading to an increase in the rate of transcription.

In both these cases, repressors, inducers and activators can be thought of as input signals into the cell. These signals are capable of affecting the expression of certain genes, and are essential when building custom triggered genetic networks with prokaryotes (Michalodimitrakis and Isalan 27). The cell can be thought of as the system while the rate of transcription is the system's output. Feedback occurs through external stimuli which control the release of the "input signals" based on whether the desired result of the protein generation is achieved.

RECOMBINANT DNA TECHNOLOGY

Hampered by current technology, the mass implementation of control engineering in genetics is rather limited. Nevertheless, the medical sector has witnessed a fundamental approach to the modelling of gene networks. This approach, though still relatively basic, is very essential as it exemplifies how genetic processes can easily be modelled and understood once they are taken as isolated subsystems with well-defined stages containing inputs and outputs.

The next section dwells on giving an explanative and descriptive model of how the process of recombinant DNA technology is used create insulin producing bacteria to aid sufferers of Diabetes. Recombinant DNA technology is a process whereby genes are spliced out of donor cells and incorporated into another organism. This investigation will look at how the genetic code required for insulin was extracted from a donor and put into *Escherichia coli* bacteria in order to create insulin. The techniques described will highlight initial methods previously used by scientists to control gene expression.

As previously mentioned, the DNA molecule has a right-handed double helix shape, with about 10 nucleotide pairs per helical turn (Pierce 266). Each strand of this double helix is made up of nucleotides bound vertically and the connection between the two strands is achieved by complementary pairs of nucleotides forming chemical bonds. To understand the application of recombinant DNA to produce insulin, the following aspects will be examined: the general structure of bacterial DNA (plasmids); the application of gene control in these organisms (Recombinant DNA formed); as well as the transfer and cloning of the products to aid mass production.

DNA IN BACTERIA: PLASMID

Bacteria are unicellular organisms, without a nuclear membrane. Most bacterial genomes (full DNA sequence), consists of a circular chromosome that contains a single DNA molecule several base pairs long. Some may have multiple chromosomes while others have linear chromosomes.

As well as possessing chromosomes, most bacteria also have *Plasmids*. Plasmids are small circular DNA molecules, which may be present in many copies per cell or just one or two per cell. These plasmids do not play an essential role in the functioning of the bacterial cell but are more important in the growth and life cycle of the cell. Some plasmids promote mating while some are used for self-defence by the bacterial cell. In general, they are vital to genetic engineering because they replicate independently from the bacterial chromosome, and unlike the complex shape of DNA in multi-cellular organisms, they have a simple circular structure.

This circular structure can be more efficiently manipulated, preventing harm to the organism (Alberts et al. 540).

Figure 3-1: A diagram of Bacterial Cell showing the circular DNA (plasmid with maker gene in orange) as well as long circular chromosome (APBI Resources for Schools 1).

INSULIN PRODUCTION IN BACTERIA USING RECOMBINANT DNA TECHNOLOGY

The process of engineering the host bacteria to make produce insulin can be divided into 4 broad stages.

STAGE 1: ISOLATING AND TREATING THE DESIRED GENE

The aim of the first stage is to isolate the segment of gene code for the particular product required and to "cut" it out. To do this, the required gene is acquired from a donor organism. It is then treated with various restriction enzymes called *Restriction Endonucleases* (Pierce 513). These enzymes serve specifically to cut the DNA at particular points thereby removing the particular gene section; in this case, the one which codes for the production of insulin. In addition, small overlaps of DNA are left sticking out at each end of the gene fragment with particular nucleotide code sequences. These are called "sticky ends" and serve a particular purpose which will be explained later (See Figure 3-2).

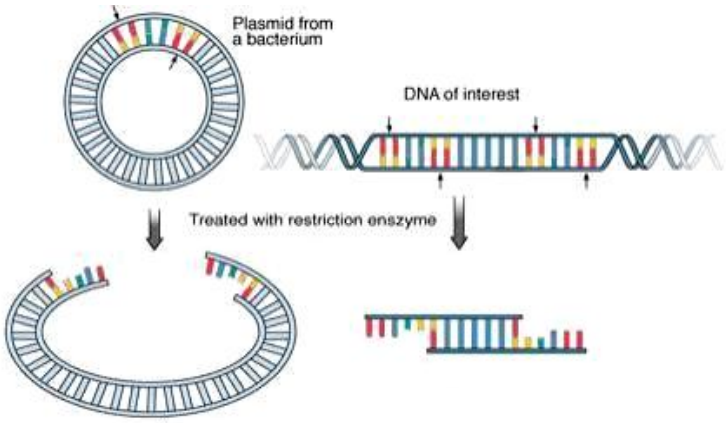


Figure 3-2: Gene Fragment (Biology Active Learner 1).

STAGE 2: IDENTIFYING AND TREATING THE PLASMID

In this stage, the vector plasmid is identified by microscopy or otherwise and isolated accordingly from the bacteria cell. The plasmid is called a vector, as it is the molecule that will carry the desired gene fragment back into the bacterial cell. The circular vector plasmid is opened up using the same restriction enzymes that were used to isolate the gene fragment. This

is done to ensure that the sticky ends created in the vector plasmid, as well as the gene fragment, overlap and are complementary to each other.

STAGES 3 & 4: COMBINATION AND TREATMENT WITH ENZYME

In stage 3, the vector and the fragment are exposed to the necessary physical and chemical conditions, such as temperature, pH, moisture, so that the complementary base pairs can recombine.

In stage 4, the product from stage 3 is then treated with another enzyme known as *DNA Ligase*. This acts as a special adhesive so that the gene fragment and vector plasmid become covalently bound together vertically (Alberts et al. 540-41).

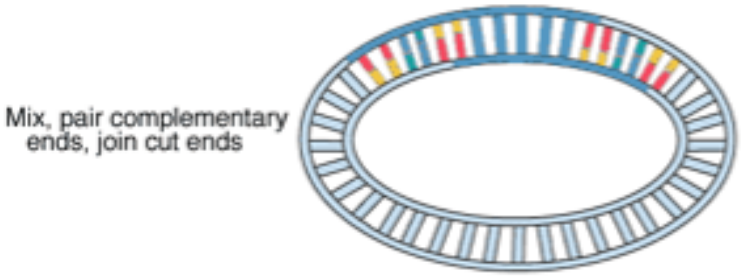


Figure 3-3: Recombinant vector plasmid carrying gene of interest (Biology Active Learner 1).

Finally, the recombinant vector plasmid is reintroduced into the bacteria cell in a process called Transformation (Pierce 520).

Though recombinant DNA technology, as described above, has proven to be success in mass production, a small proportion of cells tend not take up the recombinant plasmid. To optimise the process, a mechanism must be implemented to eliminate these unwanted cells.

To solve this problem, an additional gene, which builds resistance against an anti-biotic, is spliced alongside the desired gene. By treating the culture of bacteria with this anti-biotic following the recombination process, the undesired cells can be eliminated. Only the bacteria which have successfully accepted the plasmid will survive. The resultant culture is then incubated and the cells reproduce sporadically.

Due to the short life cycle of bacteria cells, such as *E. coli*, the recombinant plasmid is quickly replicated and passed down to offspring. When done on an industrial scale, this results in the low cost and high volume production of insulin.

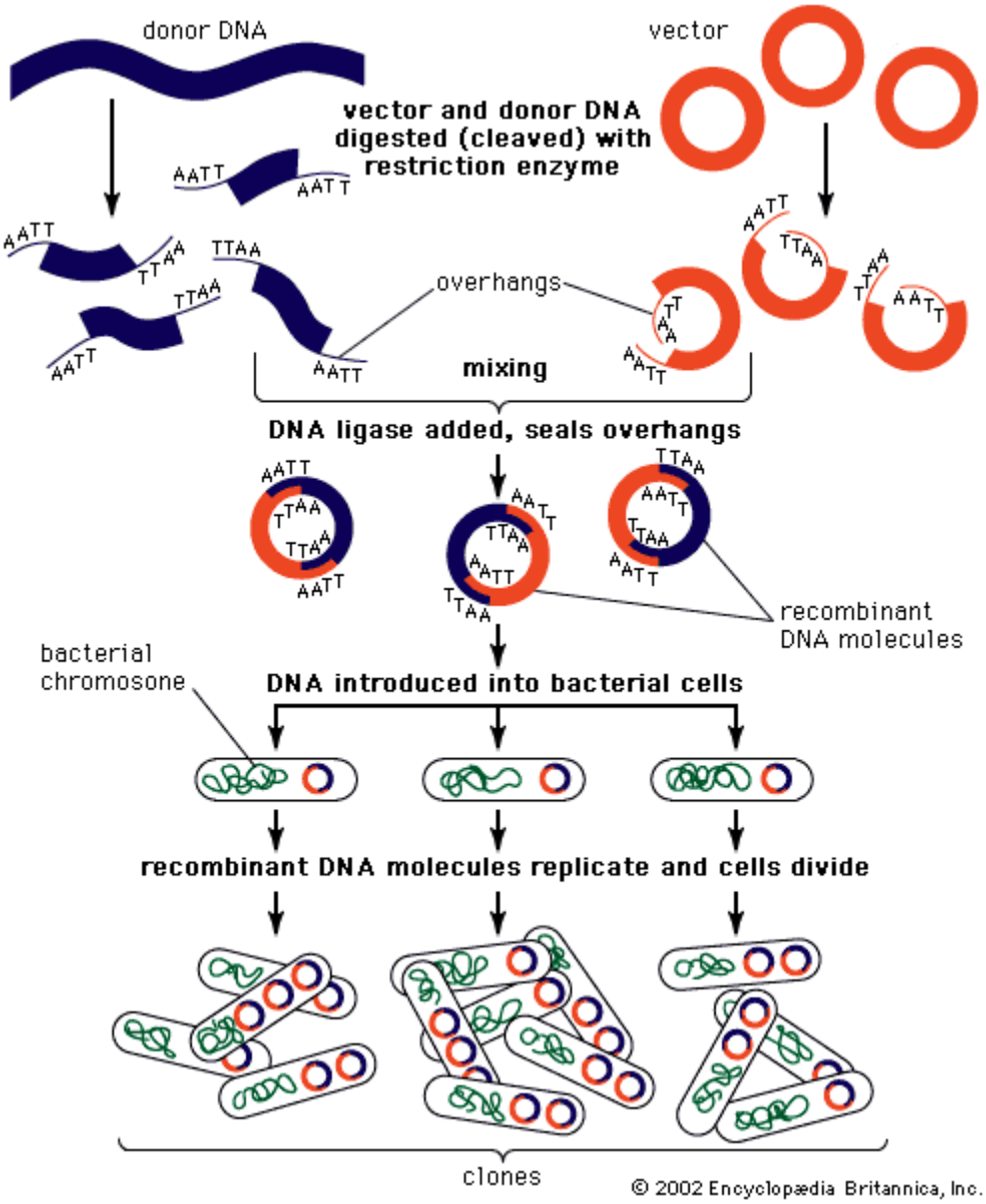


Figure 3-4: A summary of recombinant DNA technology

In the figure above, we can see recombinant DNA technology in one continuous process. The insulin created from this method is the same as those produced naturally.

MOVING BEYOND RECOMBINANT DNA TECHNOLOGY

Though recombinant DNA technology is a relatively simple and popular process, when faced with synthesizing organisms with more complex DNA, the 'copy-paste' and 'trial-error' approaches of this technology are no longer feasible.

The major disadvantage with traditional methods is that the expression of the spliced genes needs to be manually controlled. In the case of insulin production, this is done by the introduction of a chemical inducer which triggers transcription of the insulin producing gene. This process, however, is costly, error prone, time consuming and requires constant monitoring (Andrianantoandro et al. 11-12).

To overcome the limitations of traditional genetic engineering, the current generation of synthetic genetic circuits aims to implement control systems within the cells themselves. This process involves the design and construction of independent genetic decision-making circuits that may be adapted to different applications and reused. Drawing insights from engineering, a modular approach to the problem, similar to modelling a control system, is thought to be the most promising path towards realizing this goal (Hartwell et al. C48-C49).

Though biological systems are tightly integrated, it is possible to isolate small genetic networks from simple organisms such as prokaryotes. The aim is to create a well documented library of genetic parts (Voigt 548-57). Ideally these parts would serve as building blocks for complex genetic circuits. One setback may be that the inducers, activators, and proteins involved in the functioning of each module might easily interfere with other modules within the cell. This fact underlines the need for catalogued parts to be tested and documented thoroughly.

The key advantage to modular design is the ability to mathematically model the behaviour of individual genetic components. These models combined with computational analysis are thought to be the most rational approach to engineering custom genetic circuits (Michalodimitrakis and Isalan 34).

Currently, the most comprehensive library of genetic parts is hosted at the Massachusetts Institute of Technology in the United States of America. Catalogued components have been used to build modules such as environmental sensors, logic components, and biological clocks (Voigt 2-3,6). In the following sections a selected number of such modules will be examined, and real world applications highlighted where possible. A close parallel will be drawn to digital electronics, since genetic decision-making circuits are highly analogous to logic gates.

GENETIC LOGIC COMPONENTS

As highlighted earlier, key components in the latest synthetic genetic circuits are highly analogous to digital logic gates and are used to perform logic operations.

Using Boolean algebra, any combinational logic circuit can be implemented using two principal digital gates, NOT and AND/OR gates (De Morgan's Theorem). Thus the operation of these two gates using genetic circuits gives a toolbox to build networks that can scale up easily. Scaling is only limited by the number of interoperable DNA components available to be spliced.

The logic elements described here will rely on the concentration of DNA binding proteins and inducer molecules as inputs, often delivering proteins as an output. A high concentration of a molecule will be considered logic high, while a low concentration a logic low.

Similarly to silicon gates, genetic logic elements will show non-ideal characteristics in terms of switch-on and switch-off times. Protein/mRNA concentrations need to build up or die down before the gates can switch state. These times will likely be in the range of several hours and are dependent on their decay and transcription/translation rates of their relevant proteins.

The delays cannot be completely eliminated due to the chemical nature of the process. However, they can be leveraged to implement novel circuits.

BASIC LOGIC ELEMENTS: THE NOT GATE

The NOT gate is the most basic genetic logic element. It is based on the control of gene expression at the transcription level.

As formerly discussed, the first stage of gene expression involves the creation of an mRNA molecule. This molecule plays a key role linking transcription and translation. If the formation of mRNA were inhibited, no protein would result. The most common way of preventing the creation of mRNA is to impede the transcription phase. As previously described, this is achieved by introducing a repressor protein which binds itself to the promoter of the gene being controlled. This protein prevents RNAP from attaching, and hence copying, DNA segments into an mRNA transcript.

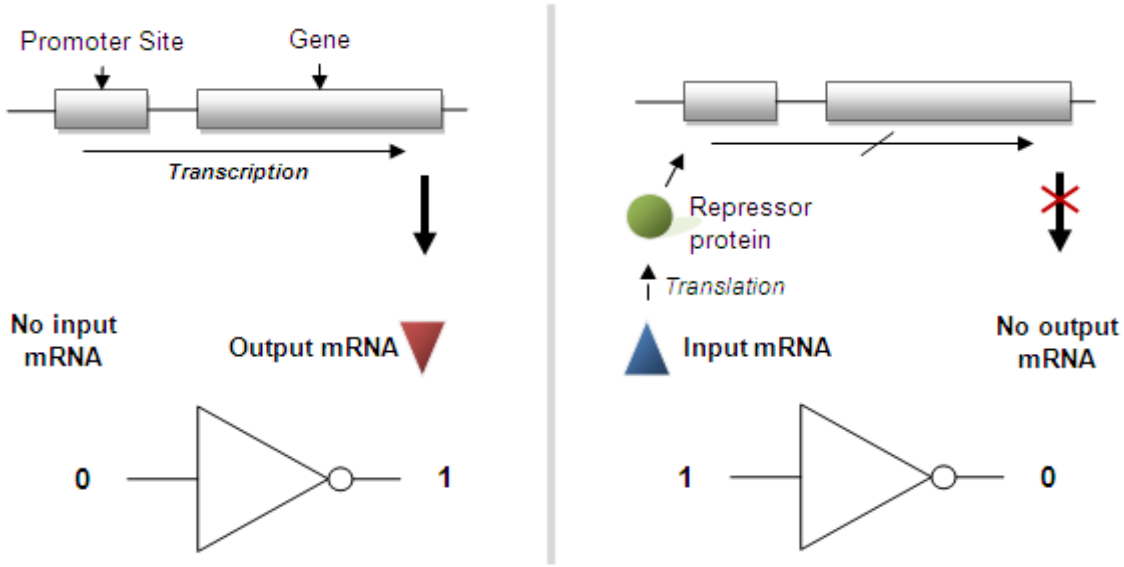


Figure 4-1: Internals of a genetic NOT gate

The repressor protein is not necessarily introduced externally. It can also be the product from another gene being expressed, that is to say, having its own input mRNA. This control system can thus be thought of as consisting of concentrations of mRNAs. If a high input concentration of the repressor mRNA is present, then the repressor protein results, and the transcription of the gene being controlled is prevented. This mechanism can be modelled using a digital NOT gate.

As shown in Figure 4-1, the absence of repressor mRNA is interpreted as a LOW input signal. This leads to the unimpeded transcription of the controlled gene, and hence a high output concentration of its mRNA, seen above as logic HIGH. When an input repressor mRNA is present, the cell will translate it and produce the repressor protein. This repressor will bind to the promoter of the gene being controlled, preventing it from synthesizing mRNA at the output. This is modelled as a LOW output signal (Weiss et al. 49-52).

These principles can be extended to build other types of gates which rely on mRNA concentrations for their inputs and outputs. For instance a NAND gate can be built using two NOT gates which synthesise the same protein at their outputs. This effectively performs an OR operation on these outputs, leading to a functional NAND gate (Weiss et al. 52). However, these gates, which rely heavily on repressor proteins, are limited to inter-cellular operations. In order to build more complex multi-cellular genetic networks, there is a need to implement methods for intercellular communication.

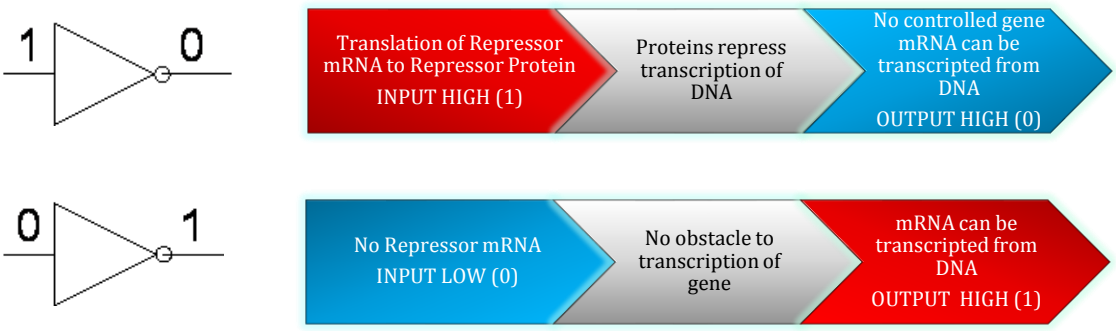


Figure 4-2: Overview of genetic NOT gates

RING OSCILLATOR WITH CASCADED NOT GATES

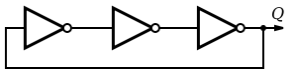


Figure 4-3: A Ring Oscillator

The genetic ring oscillator follows the same principle as used in electronics, with an odd number of NOT gates cascaded in series, and the output of the last NOT gate being fed back as input to the first. The oscillation is due to the propagation delay of the signal through the NOT gates, and is dependent on the gate switching time.

Genetic ring oscillators demonstrated in laboratory environments often have 3 cascaded NOT gates. The protein output of the last NOT gate is a repressor for the first NOT gate. The output of the former also leads to the creation of a protein that fluoresces so that the state can be determined. The oscillation period of this circuit is in the region of several hours, with a high variation due to the uncertainty in gate switching (Michalodimitrakis and Isalan 28). The apparent noise seen in a culture of ring oscillators is a result of the uncoordinated expression of genes in these cells.

Experiments done at the University of Princeton involve engineering this same process, albeit without the abstraction of genetic gates (Elowitz and Leibler 335-38). Time-lapse photos of the engineered bacterial culture are shown in Figure 4-4.

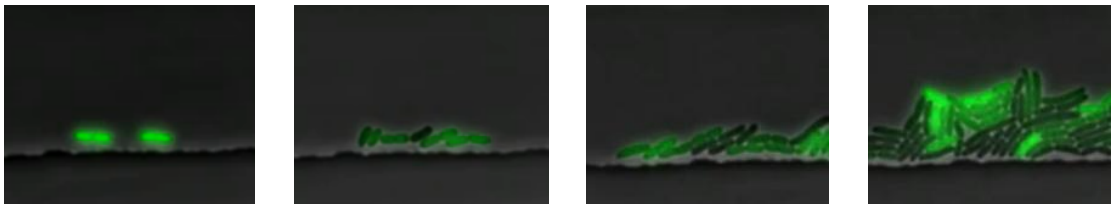


Figure 4-4: Genetic ring oscillator at 5, 200, 400 and 550 minutes

INTER-CELLULAR COMMUNICATION

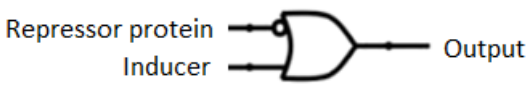
Controlling the expression of genes in single cells is limited by the complexity of the tasks which can be performed. More complex tasks will require a higher level of coordination between engineered cells. A large number of cells synchronously executing a single task can also reduce the uncertainty posed by the failure of individual cells (Andrianantoandro et al. 10).The latest research into inter-cellular signalling has proven that even the simplest of organisms can coordinate their internal processes using small chemical messengers called inducers (Camilli and Bassler 1113-16).

IMPLIES and AND genetic gates have been designed to receive signals from beyond the cell membrane. Externally originating signals can be received by the cell in the form of inducer molecules. A genetic component can be implemented within transmitter cells to encode for the inducer molecules needed.

Compelling demonstrations of this technology have been showcased through recent research. Internal design of base communication systems and basic examples of cells co-ordinating for time synchronicity are discussed in the next sections.

IMPLIES GATES

The IMPLIES gate has two inputs, a repressor protein and an inducer molecule, and produces a single protein output. When the inducer input is LOW, the behaviour of this network is identical to that of a NOT gate. Increases in the inducer concentration outside the cell, considered a HIGH inducer input signal, will cause the inducer molecules to diffuse through the cell membrane. Repressors will preferably bind to the inducer molecules entering the cell, leaving the promoter site of the output gene unblocked. RNAP can then transcribe this gene, leading to the synthesis of its protein product: a HIGH output (Weiss et al. 53). The truth table for this gate can be seen in Table 4-1.



Repressor	Inducer	Output
0	0	1
0	1	1
1	0	0
1	1	1

Table 4-1: Truth table of an IMPLIES gate

Note that the output protein of this module can be a repressor/activator for another gene or an inducer for further communication.

AND GATES

Similar to an IMPLIES gate, the AND gate has two inputs and produces a single protein output. However, instead of a repressor input, the AND gate has an activator input (Weiss et al. 54).

Activators attach to the promoter site of their target gene and assist RNAP in attaching to the gene and to then transcribe it. The controlled gene's promoter site in an AND gate has a low affinity for RNAP. Thus in the absence of a functional activator, the gene will not be expressed and the protein output will be LOW.

The activator proteins at the input of the AND gate are chosen to be non-functional in the absence of the inducer molecules. Thus only a HIGH input from both the inducer and the activator will result in a sufficient rate of transcription of the output gene, and hence lead to a HIGH output. In all other cases the output protein would be LOW in concentration.

As with the IMPLES gate, inducer molecules may either be introduced externally, or produced internally by the expression of another gene.

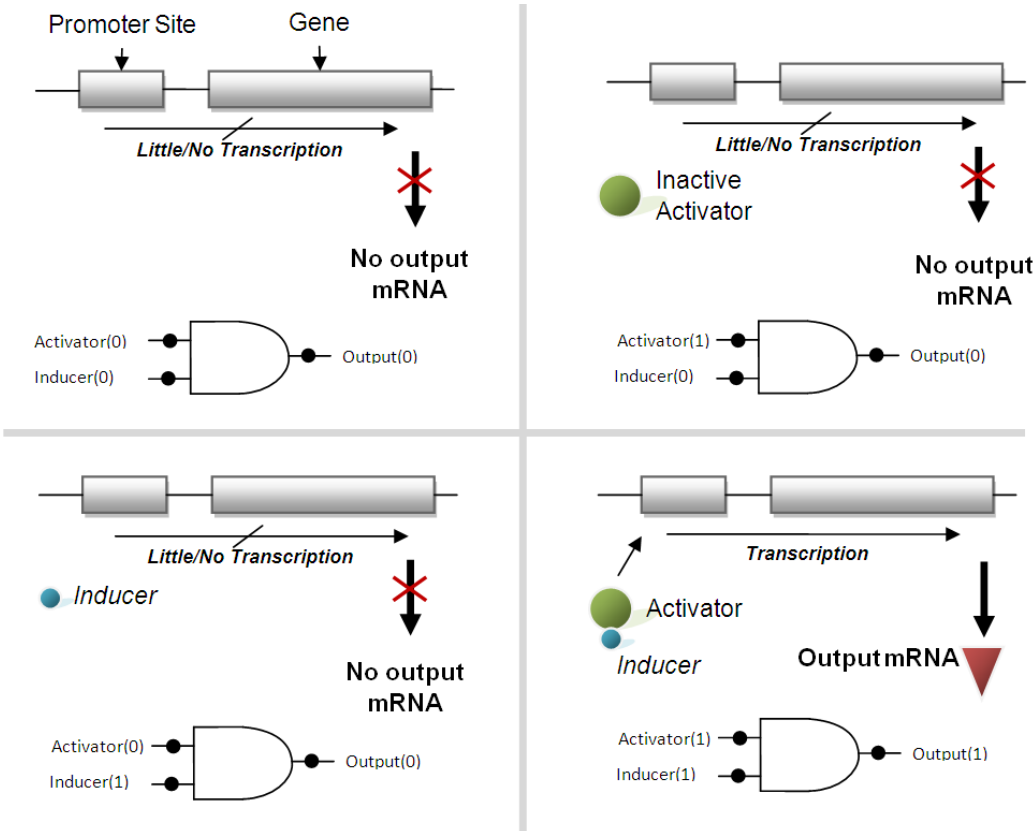


Figure4-5: Internals of a genetic AND gate

A SIMPLE EXAMPLE OF *AND* AND *IMPLIES* GATES

A marine bacterial species named *Vibrio fischeri* has the ability to detect its cell culture density. When its concentrations are greater than 10^{10} cells per litre, this culture produces bioluminescence. The bacteria makes this decision through the synthesis and detection of a chemical inducer called *V. Fischeri auto-inducer* (VAI) (Fuqua, Winans, and Greenberg 1).

The inducer molecule is able to diffuse through cellular membranes and sea water, creating chemical gradients. At low cell densities, the inducer will diffuse through the medium, resulting in a low concentration inside the cell. However, at high enough cell densities, the inducer concentrations inside the cell will be sufficient to trigger the synthesis of the luminescent protein. This cell density detection mechanism has been termed Quorum Sensing.

By isolating the genetic components responsible for quorum sensing in *V. Fischeri*, it is possible to splice its genes into *E. Coli* bacteria cells and build versions of the aforementioned AND and IMPLIES gates (Knight Jr. and Weiss 1-16).

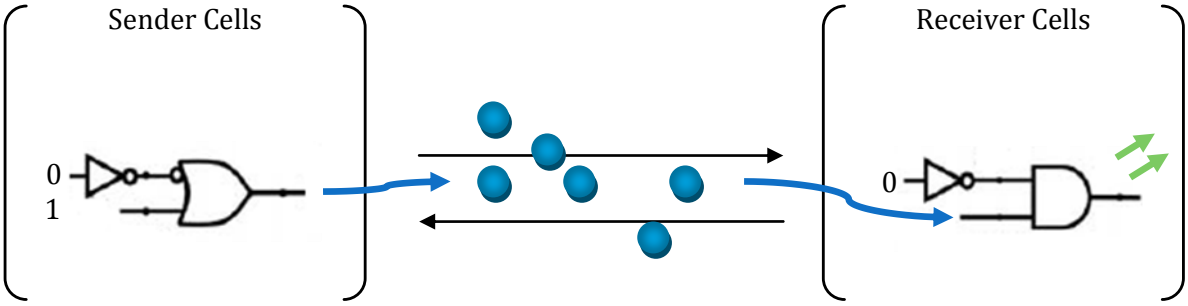


Figure 4-6: Schematic of a simple inter-cellular communication system

The NOT gate in the receiver results in the continuous expression of the repressor proteins for the IMPLIES gate. The second input can be controlled by another cellular network. When this second input is HIGH, a protein is created which catalyses the synthesis of the inducer VAI. These inducers diffuse out through the sender's cell membrane and then spread into the receiver cells.

In the receiver cells, the NOT gate synthesises non-functional Activators for the AND gate. As soon as the inducer input to the AND gate is HIGH, inducers will attach to the activators, rendering them functional. This leads to the expression of a fluorescent protein at the output.

When a culture of sender cells are left surrounded by a culture of receiver cells, the response to communication can be seen through visual identification of fluorescence. This system provides a basic real-world example of intercellular communication using reusable logic blocks (Weiss 75).

CELL CO-ORDINATION IN THE TIME DOMAIN: A BIOLOGICAL CLOCK

Time coordination of cell operations in a culture is crucial to streamlining and optimising industrial processes that depend on genetic engineering. The core logic units and communication mechanisms discussed in previous sections have been applied to synchronise a genetic oscillator across a whole culture of bacteria. This is thought to be the first step towards a viable synthetic biological clock. Intercellular communication is facilitated by the genetic parts isolated from the quorum sensing mechanism of *V. fischeri*.

A single inducer molecule can trigger the expression of three genes. These three network paths constitute of a fluorescent output, a negative feedback loop and a positive feedback loop (Fussenegger 301-02). The output path is simply wired to a promoter which codes for a green fluorescent protein which is used to detect the state. The negative feedback path codes for an enzyme which accelerates the decay of the inducer molecules themselves. The positive feedback loop codes for an enzyme that leads to the creation of inducers themselves. By balancing the strengths of each control loop an oscillation can be created.

Since inducers can diffuse through the cell membrane when sufficient concentrations exist, they form a synchronising signal. This results in the entire group of cells flashing in unison. However, due to the diffusion speed of inducer molecules in the medium, the fluorescent proteins flash in waves.

This technology paves the way towards a whole new class of genetic control applications. An example would be insulin producing bacteria that are embedded within the human body, and releases doses of insulin at pre-programmed time intervals.

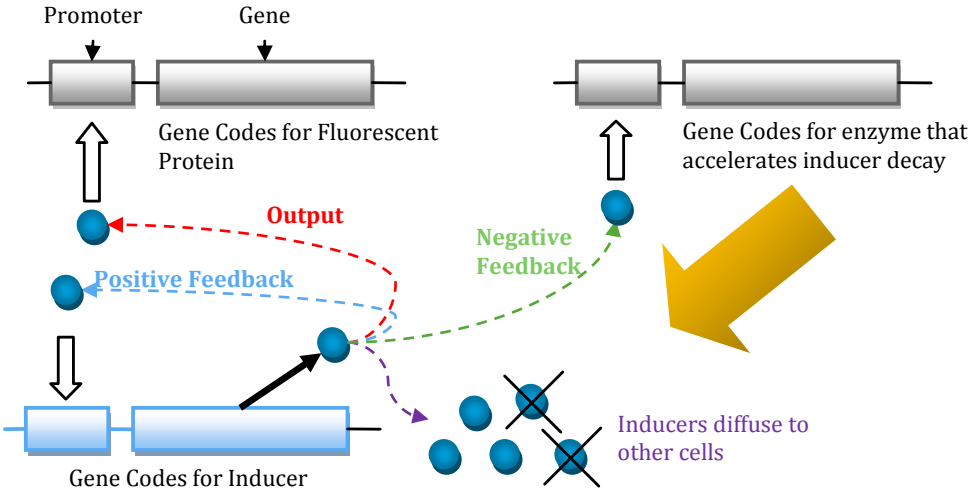


Figure 4-7: Overview of the genetic oscillatory system

PERFORMANCE ENHANCMENT IN GENETIC NETWORKS

The existence of various dynamic factors and variables can make controlling the performance of genetic networks quite challenging. Changing the network construction is a possible solution, but this process is very time-consuming, error prone, and susceptible to conflicts due to module incompatibilities. Expanding networks to include multiple modules may also complicate and prevent the desired performance being achieved. Seemingly, the process of building synthetic genetic networks can become a very complex engineering system. However, there is one important property of genetic networks which may be potentially be utilized: Evolution (Yokobayashi, Weiss, and Arnold 16587-91).

Genetic networks are able to evolve under environmental conditions. This natural process can be guided using a technique known as Directed Evolution. Directed evolution involves cells being bred in an artificially designed environment where their survival depends on how well they perform logical operations. To achieve optimal results, network parameters are randomly set on different cell cultures. To observe the response of the cell cultures, the output of the logical network will produce florescent proteins.

After various tests, the most successful cell colonies are selected. These samples are further analysed and the process is repeated various times until the minimum required standard is achieved for each cell culture. The surviving colonies are then implemented as prototypes. Interestingly, some successful colonies have been seen to develop structures initially thought to be non-functional, yet these cells have evolved to adopt certain structures and traits making them fully operational (Yokobayashi, Weiss, and Arnold 16587-91).

Nevertheless, it must be noted that this technique also has its limitations. While evolution may self-optimise synthetic genetic networks, it may also render them inoperable. It must be noted that cells do their best to sustain survival; however, synthetic networks can be too heavy a burden on the cells metabolism or interfere with their natural processes. This leads to high probabilities of cells dying or networks being deactivated through evolution in subsequent generations.

THE FUTURE OF SYNTHETIC BIOLOGY

Undoubtedly, abstraction of genetic networks as modules was a very profound step toward making synthetic biology operational for medical use. This abstraction allowed researchers to focus on the overall function of the modules while ignoring the low-Level chemical interactions. This approach also increased the flexibility of the overall network.

CURRENT ISSUES WITH MODULARITY APPROACH

Out of the 5000 different building blocks commercially available, only a very small number are fully functional. Lack of knowledge regarding the properties of each individual gate as well as their high susceptibility to noise renders their behaviour very unpredictable. Furthermore, unlike digital gates, genetic gates cannot be easily debugged. Therefore, connecting multiple modules together can often lead to incompatibilities and result in a non-functional network. In addition, quite often, the externally introduced genes are subject to interference from the existing natural genes inside the cell, causing the cell to die (Kwok 289).

CELL CHASSIS: A BLANK PLATFORM

One way to solve interference between the inserted gene and the natural internal genes is to create Platform Cells. Platform cells contain a minimum quantity of internal genes allowing them to keep themselves alive, while, at the same time, creating the natural cell's gene artificially. In this method, genes are created from scratch as opposed to using already existing libraries, meaning that there is full control over every little part of the gene's function. By using these blank platform cells, called Cell Chassis, not only is there a better chance of getting the gate to work as well as having greater control over the properties of the gate, but there is also less risk of cell termination (Andrianantoandro et al. 9).

Though the method of platform cells is very promising, it is currently not easy to implant the entire gene required to make a fully functional cell. In addition, it is usually difficult to create the chassis cells with a minimum amount of internal genes. Furthermore, this approach does not eliminate the problem of incompatibility between different networks. Upon connection, the overall configuration changes and a new network with unique characteristics is created. However, the main issue with this method is that it is currently very expensive and labour intensive to apply.

Current research focuses on creating and using bacteria to speed up the process of inserting the gene into cells. E-coli cells are used as "assemblers" and are equipped with certain enzymes that enable them to "cut" and "paste" different segments of DNA in the cell. Their function is

somewhat similar to viruses and they have the information to assemble the entire gene from disjoint segments (Kwok 288).

SELECTIVE DNA TRANSLATION

In an alternative approach, the cell's internal DNA is distinguished from the DNA related to the functioning of the gate. Enzymes are employed in this method, to make sure that the DNA sequence, which is being copied, does not exist on the cell's natural gene. This technique effectively reduces the chance of clash between the cell's internal functions and its role to create the required proteins. In a similar approach, researchers try to create physical isolation using a "membrane" between the cell's internal genes and the genes related to its function as a gate (Kwok 290).

RIBOSWITCHES

Using gates and proteins to suppress the transcription stage is inefficient as cells must create a repressor protein which will only be used once. This protein will decay away over time and until it has done so completely, a new one will not be produced. An alternative method is to use Riboswitches which are sections of mRNA that can directly regulate the behaviour of mRNA and hence control the gene's activities (Tucker and Breaker 342). These sections on the mRNA are highly sensitive to metabolites, which are small molecules, usually proteins, which connect to specific regions of mRNA and control these Riboswitches.

Riboswitches are made of two parts: an aptamer and an expression platform. The aptamer is in charge of selectively binding to the corresponding metabolites, while the expression platform, which, depending on the metabolite that the aptamer is bound to, changes the expression of genes. The aptamer's structure cannot be easily modified, however, it is possible to change the structure of expression platform and acquire different functionalities. By externally controlling metabolites, it is theoretically possible to control the gene activities. The technology of controlling riboswitches is very recent and hypothetical. However, if applied in practice, it would be feasible for large-scale situations (Adaya 194).

METABOLIC CONTROL

When cells are used as digital gates, a problem usually encountered is to keep up with the high demand of protein that is needed for their digital operation. In an effort to overcome this problem, a control system is employed to constantly measure the metabolic activity of cells and change the demand for the output protein as a function of the internal metabolic activities of the cell. This approach, known as Metabolic Engineering, concentrates on changing the internal

regulatory genes and adjusting the protein output demand with the metabolic process of the cell. As a result, cells will produce the required proteins more efficiently. This approach leads to less resultant toxic waste from the output protein, and the cells are less likely to die as a result of being overloaded with a high demand of protein production (Farmer and Liao 534).

The feasibility of this method was tested during an experiment using *E.Coli*. The aim of the experiment was to improve the production of a compound used for cancer treatment termed *Lycopene*. The level of glucose gave a measure of metabolic activities in the cell. A type of phosphate, called ACP, was employed to measure the level of glucose. This phosphate regulated the expression of certain genes that controlled the amount of a promoter. The promoter regulated the genes on DNA which were in charge of controlling the metabolic activities of the cell. This control loop aimed to dynamically control the flux of glucose according to the metabolic state of the cell, in order to produce the least amount of toxic waste and to maximise the production of Lycopene (Farmer and Liao 535).

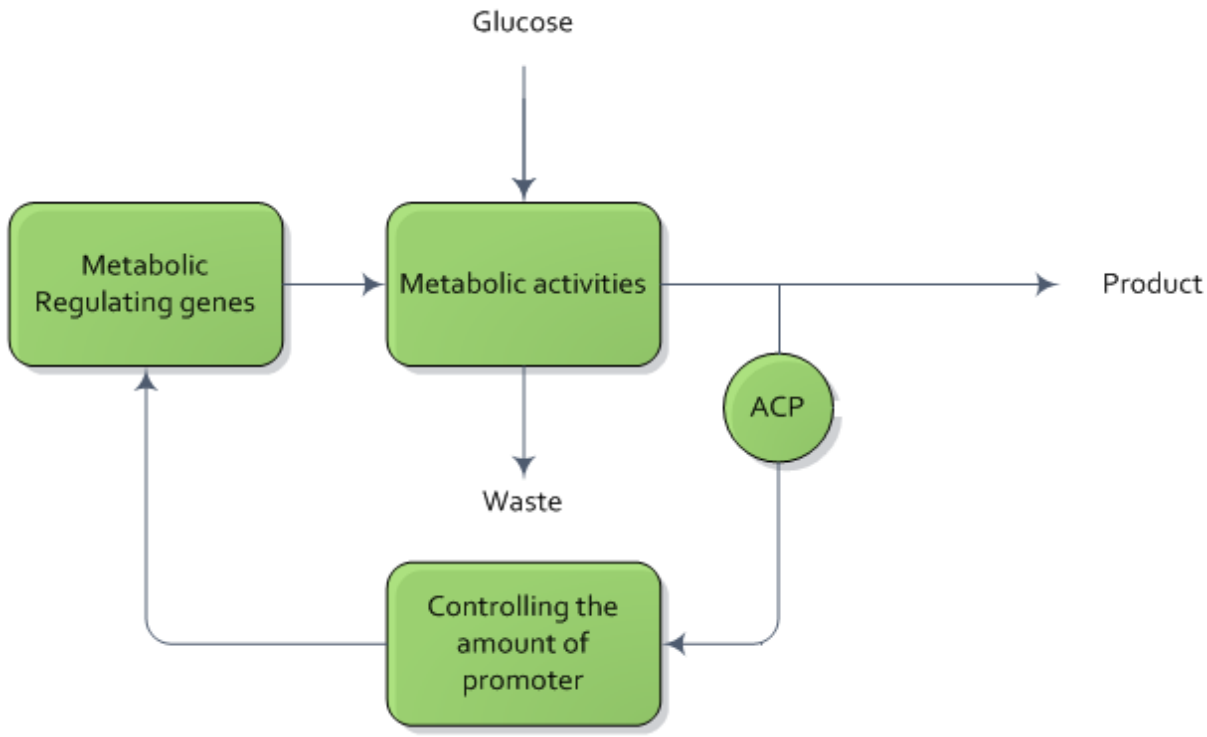


Figure 5-1: Overview of the Lycopene production system

SOCIAL AND ETHICAL ISSUES

As the ability of modifying gene structures continues developing, there is growing concern as to where this technology may lead. Some argue that this technology has and will continue saving many lives while others are worried that it is becoming increasingly easier to manipulate genetic structures, opening a window for people with malicious intentions. In fact, both claims are equally valid.

Over the past decade the speed and costs of sequencing have advanced faster than computers. The ability to create new complex organisms has virtually become a reality. Instruction sets implemented into foreign cells will one day enable scientists to assemble an artificial life form. This soon to be existent process has been described as "software that creates its own hardware" by various scientists (Krauss 1).

While this potential technique promises many benefits, it also raises many concerns. The extent to which these manually synthesized organisms can be contained safely in laboratories is an important hazard. Given that this technology is still very young and in testing stages, any defect or irregularities found in offspring cells could lead to potential health risks if released into the environment. The increasing capability for scientists to create DNA sequences may also encourage unethical uses of this practice causing potential proliferation of contaminations or distortions in organisms' DNA (Krauss 1).

If this technology were used to release doses of insulin at pre-programmed time intervals for humans with diabetes, as previously discussed, the issue of reliability must also be taken into account. Future mass implementation of synthetic biology would require countless rigorous tests to ensure reliability and compatibility under all conditions. Though, at present, every effort has been made to increase reliability, such as digital gate abstraction, there are still many dynamic factors and variables in biological organisms which cannot be accounted for and side effects are almost inevitable.

Furthermore, the concept of artificially creating living organisms raises many concerns with philosophers and the religious community. Evolution and heredity have always been seen as natural processes, however, if synthetic biology were to prevail, questions of what is now considered natural and whether it is correct would surface. Interference with natural systems may pose many disputes, but it could also save many lives and find the cures to widespread diseases such as HIV/AIDS and cancer. Although its effects on future generations and the environment are yet to be seen, social and legal factors will be play an important role in determining the future on the control of gene expression in the field of synthetic biology.

CONCLUSION

Undoubtedly, control of gene expression is a newly born field revolutionizing the worlds of science and engineering. Future outcomes of this technology are extremely promising with new possibilities emerging every year. Nearly every disease commences at the level of gene expression. The ability to potentially control which genes get expressed in individual cells offers future optimism to cure almost any disease (UT Southwestern Medical Center 1).

This report has investigated different techniques used both currently and in the past, to control the expression of genes in relatively simple cells. Organisms of increasing complexity have set limitations on the aforementioned methods, leading to the extraction of cellular systems into modules. This isolation has enabled the construction of large genetic networks showing similar resemblances to digital logic gates. While these networks have shown competencies for future development, they take very long to evolve and very few have demonstrated full functionality.

As a result, a new method for control of gene expression has emerged. Rather than building genetic networks using the analogy of digital gates, new approaches propose to use natural processes within the cell to maximize efficiency and reduced interference during gene expression. While these developments still remain in the testing phases, it promises great advantages over the modular approach.

Regardless of the ultimate method used, controlling gene expression has confirmed man-made biological machines are an achievable reality, and as such, it has created a new level of engineering. It is not far from the imagination that in the next few decades the next generation of genetic computers, nano-robots made from proteins, will be designed to cure cancer and clear infections (Ball). Perhaps this new level of engineering will overshadow the environment in triggering evolution for future generations with the criteria for survival of the fittest changing. Nonetheless, whichever developments result from synthetic biology, there will be one key factor determining the extent to which they transpire: technology.

BIBLIOGRAPHY

Adaya, N. Cohen. "Integration of RNA Search Methods for Identifying Novel Riboswitch Patterns in Eukaryotes". 2005.V 193-95.

Alberts, Bruce, et al. Molecular Biology of the Cell: Reference Edition. 5th ed. 2007.

Andrianantoandro, Ernesto, et al. "Synthetic biology: new engineering rules for an emerging discipline." Mol Syst Biol 2 (2006).

APBI Resources for Schools. Diagram of Bacterial Cell. 2010. 4-12-2009.

Ball, Philip. "Synthetic biology for nanotechnology." (2005).

Biology Active Learner. 6th. 2002. Thomson Learning, Inc. 17-12-2009.

Biotol, Benjamin M, B. C. Currell, and R. C. Dam-Mieras. Genome Management in Prokaryotes. 1st ed. Butterworth-Heinemann, 1993, 160-83.

Camilli, Andrew and Bonnie L. Bassler. "Bacterial Small-Molecule Signaling Pathways." Science 311.5764 (2006): 1113-16.

Doyle, John, Francis Bruce, and Allen Tannenbaum. Feedback Control Theory. Macmillan Publishing Co., 1992.

Elowitz, Michael B. and Stanislas Leibler. "A synthetic oscillatory network of transcriptional regulators." Nature 403.6767 (2000): 335-38.

Farmer, William R. and James C. Liao. "Improving lycopene production in Escherichia coli by engineering metabolic control." Nat Biotech 18.5 (2000): 533-37.

Fuqua, Claiborne W., Stephen C. Winans, and E. Peter Greenberg. "Quorum Sensing in Bacteria: the LuxR-LuxI Family of Cell Density-Responsive Transcriptional Regulators." Journal of Bacteriology.Jan. 1994 (2010).

Fussenegger, Martin. "Synthetic biology: Synchronized bacterial clocks." Nature 463.7279 (2010): 301-02.

Goodchild, Lucy. Heart rhythm gene revealed in new research.Imperial College London, 2010.

Hartwell, Leland H., et al. "From molecular to modular cell biology." Nature.

Knight Jr., Thomas F. and Ron Weiss. "Engineered Communications for Microbial Robotics." DNA Computing. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010. 1-16.

Kwok, Roberta. "Five hard truths for synthetic biology." Nature 2010: 288-90.

Latchman, David. Gene Regulation: A Eukaryotic Perspective. 4th ed. Taylor & Francis, 2005.

Lewin, Benjamin M. The Genetic Code. Wiley, 1971.

Michalodimitrakis, Konstantinos and Mark Isalan. "Engineering prokaryotic gene circuits." FEMS Microbiology Reviews 33 (2009): 27-37.

National Library of Medicine. Cells and DNA : What is DNA?Genetics Home Reference, 2010.

Pierce, Benjamin A. Genetics: A Conceptual Approach. 2nd ed. W. H. Freeman and Company, 2002.

Tucker, Brian J. and Ronald R. Breaker. "Riboswitches as versatile gene control elements." Current Opinion in Structural Biology 15.3 (2005): 342-48.

Twyman, Richard. Gene Expression.Wellcome Trust, 2003.

UT Southwestern Medical Center. Gene Silencing Technique Can Theoretically Cure Any Disease.RxPG News, 2005.

Voigt, Christopher A. "Genetic parts to program bacteria." Current Opinion in Biotechnology 17.5 (2006): 548-57.

Weiss, Ron. "Cellular Computation and Communication Using Engineered Genetic Regulatory Networks." Diss. 2001.

Weiss, Ron, et al. "Genetic circuit building blocks for cellular computation, communications, and signal processing." Natural computing 2.1 (2003): 47-84.

Yokobayashi, Yohei, Ron Weiss, and Frances H. Arnold. "Directed evolution of a genetic circuit." Proceedings of the National Academy of Sciences of the United States of America 99.26 (2002): 16587-91.